There are often no suitable cell-surface markers to use in isolating single cells for study, and even when there are, a small number of single cells are not sufficient to capture the range of natural variation in gene expression. What is needed is a method of preparing cDNA libraries which can be used to analyze gene expression in a plurality of single cells.

SEQUENCE LISTING

The present application is being filed along with a Sequence Listing in electronic format. The Sequence Listing is provided as a file entitled IP-1245-PCT_SequenceListing.txt, created April 28, 2015, which is 1 Kb in size. The information in the electronic format of the Sequence Listing is incorporated herein by reference in its entirety.

SUMMARY

Presented herein are methods and compositions for multiplexed single cell gene expression analysis. Some methods and compositions include the use of droplets and/or beads bearing unique barcodes such as unique molecular barcodes (UMI).

Accordingly, one example presented herein is a method of preparing a cDNA library from a plurality of single cells, the method comprising the steps of: releasing mRNA from each single cell to provide a plurality of individual mRNA samples, wherein the mRNA in each individual mRNA sample is from a single cell; synthesizing a first strand of cDNA from the mRNA in each individual mRNA sample with a first strand synthesis primer and incorporating a tag into the cDNA to provide a plurality of tagged cDNA samples, wherein the cDNA in each tagged cDNA sample is complementary to mRNA from a single cell. In one example, the tag comprises a cell-specific identifier sequence and a unique molecular identifier (UMI) sequence. In some examples, the tag comprises a cell-specific identifier sequence without the UMI. The method further comprises pooling the tagged cDNA samples; optionally amplifying the pooled cDNA samples to generate a cDNA library comprising double-stranded cDNA; and performing a tagmentation reaction to simultaneously cleave each cDNA and incorporate an adapter into each strand of the cDNA, thereby generating a plurality of tagged cDNA fragments. In some examples, sufficient number of single

cells is present, the amplification of cDNA can be avoided. The second strand of cDNA is synthesized using a template switching oligonucleotide primer (TSO primer), followed by symmetric Nextera.

In one aspect, the present invention provides a method of preparing a cDNA library from a plurality of single cells, the method comprising the steps of: releasing mRNA from each single cell to provide a plurality of individual mRNA samples, wherein the mRNA in each individual mRNA sample is from a single cell; synthesizing a first strand of cDNA from the mRNA in each individual mRNA sample with a first strand synthesis primer and incorporating a tag into the cDNA to provide a plurality of tagged cDNA samples, wherein the cDNA in each tagged cDNA sample is complementary to mRNA from a single cell, and wherein the tag comprises a cell-specific identifier sequence and optionally a unique molecular identifier (UMI) sequence; pooling the tagged cDNA samples; optionally amplifying the pooled cDNA samples to generate a cDNA library comprising double-stranded cDNA; and performing a tagmentation reaction to simultaneously cleave each cDNA and incorporate an adapter into each strand of the cDNA, thereby generating a plurality of tagged cDNA fragments.

In certain examples, the method further comprises amplifying the tagged cDNA fragments to generate amplified tagged cDNA fragments. In some examples, amplifying comprises adding additional sequence to the 5' end of the amplification products.

In some examples, the additional sequence comprises primer binding sequence for amplification on a solid support. In some examples, the additional sequence comprises additional index sequences.

In certain examples, the method further comprises amplifying the amplified tagged cDNA fragments on a solid support.

In certain examples, the method further comprises sequencing the amplification products on the solid support.

In some examples, the tagmentation reaction comprises contacting the double-stranded cDNA with a transposase mixture comprising adapter sequences that are not found in the first strand synthesis primer.

In some examples, the transposase mixture consists essentially of transposomes having one type of adapter sequence.

In certain examples, the method further comprises sequencing the tagged cDNA fragments.

In some examples, sequencing comprises 3' tag counting.

In some examples, sequencing comprises whole transcriptome analysis.

In some examples, first strand synthesis is performed using a mixture of random primers, the random primers further comprising a tag.

In some examples, the first strand synthesis primer comprises a double-stranded portion. In some examples, the first strand synthesis primer comprising a double-stranded portion further comprises a single stranded loop at one end. In some examples, the first strand synthesis primer comprises a region capable of forming a hairpin.

In some examples, the first strand synthesis primer reduces concatenation byproducts compared to a single-stranded first strand synthesis primer.

In some examples, the first strand synthesis primer comprises a region of RNA.

In some examples, the first strand synthesis primer is hybridized to a complementary oligonucleotide, thereby forming a double stranded portion.

Also presented herein is a plurality of beads, wherein each bead comprises a plurality of oligonucleotides, each oligonucleotide comprising: (a) a linker; (b) an amplification primer binding site; (c) optionally a Unique Molecular Identifier which differs for each oligonucleotide; (d) a bead-specific sequence that is the same on each oligonucleotide on the bead but is different on other beads; and (e) a capture sequence for capturing mRNAs and priming reverse transcription.

In some examples, the capture sequence comprises oligo-dT.

In some examples, each bead is in a separate droplet segregated from other beads.

The details of one or more embodiments are set forth in the accompanying drawings and the description below. Other features, objects, and advantages will be apparent from the description and drawings, and from the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is an exemplary schematic showing a schematic of cDNA synthesis using SMARTer methodology.

Fig. 2A is an exemplary schematic showing first strand synthesis according to one embodiment, primed with oligo dT that has been appended with an optional molecular barcode (UMI), a sample barcode (BC), an amplification primer binding sequence (V2.A14) and a template switch (TS) primer sequence, followed by template switching, pooling of samples and single primer PCR.

Fig. 2B is an exemplary schematic showing tagmentation of pooled amplification products with symmetric Nextera, followed by amplification using different forward (V2.B15) and reverse (V2.A14) PCR primers and subsequent paired end sequencing.

7. The method of claim 1, wherein the tagmentation reaction comprises contacting the double-stranded cDNA with a transposase mixture comprising adapter sequences that are not found in the first strand synthesis primer.

8. The method of claim 7, wherein the transposase mixture consists essentially of transposomes having one type of adapter sequence.

9. The method of claim 1, further comprising sequencing the tagged cDNA fragments.

10. The method of claim 9, wherein sequencing comprises 3' tag counting.

11. The method of claim 9, wherein sequencing comprises whole transcriptome analysis.

12. The method of claim 1, first strand synthesis is performed using a mixture of random primers, the random primers further comprising a tag.

13. The method of claim 12, wherein the first strand synthesis primer comprises a double-stranded portion.

14. The method of claim 13, wherein the first strand synthesis primer reduces concatenation byproducts compared to a single-stranded first strand synthesis primer.

15. The method of claim 13, wherein the first strand synthesis primer comprises a region capable of forming a hairpin.

16. The method of claim 13, wherein the first strand synthesis primer comprises a region of RNA.

17. The method of claim 13, wherein the first strand synthesis primer is hybridized to a complementary oligonucleotide, thereby forming a double stranded portion.